



多目的行動調停に基づく知能ロボットの行動獲得に関する研究

能島, 裕介

(Degree)

博士 (工学)

(Date of Degree)

2004-03-31

(Date of Publication)

2013-03-12

(Resource Type)

doctoral thesis

(Report Number)

甲3158

(URL)

<https://hdl.handle.net/20.500.14094/D1003158>

※ 当コンテンツは神戸大学の学術成果です。無断複製・不正使用等を禁じます。著作権法で認められている範囲内で、適切にご利用ください。



【 3 2 7 】

氏 名・(本 籍)	能島 裕介	(富山県)
博士の専攻分野の名称	博士 (工学)	
学 位 記 番 号	博い第337号	
学位授与の 要 件	学位規則第4条第1項該当	
学位授与の 日 付	平成16年3月31日	

【 学位論文題目 】

多目的行動調停に基づく知能ロボットの行動獲得に
関する研究

審 査 委 員

主 査	教 授	小島 史男
	教 授	田浦 俊春
	教 授	上原 邦昭

近年、知能ロボットの開発が盛んに行われ、生産現場から極限作業環境や生活環境など様々な環境でロボットが適用されている。極限作業環境や生活環境において、知能ロボットが考慮すべき目標（自己維持、探索、追従、採取など）が複数存在し、同時にそれらを満たす多目的な動作が知能ロボットに要求されている。この多目的な動作は、個々の目的に対応した行動出力により構成されるが、耐故障性や人間との円滑なコミュニケーションのために、一連の動作出力のギャップを緩和する滑らかさが必要とされる。また、極限作業環境や生活環境は、事前に想定できない未知な環境であり、その環境に適した行動を獲得する必要がある。さらに、獲得した行動は、新しい環境やタスクによる新しい状況に迅速に対応するために、再利用できなければならない。

本研究では、時系列観測情報を考慮した多目的行動調停に基づく知能ロボットの枠組みを提案し、様々な状況に対応するための多目的かつ滑らかな動作生成及び、それを実現する基本行動と行動調停則の学習手法を提案する。基本行動は、タスクを達成するために必要な複数の目的に対してそれぞれ構成され、行動調停則は、各基本行動に対する行動重みを更新するルールにより構成される。多目的行動調停とは、どの基本行動をどのくらい用いるかを、時系列観測情報を考慮してヒューリスティックに決定し、各基本行動の行動出力を重み付け平均で融合する手法である。行動調停則は、基本行動に対してトップダウン的に役割を与えるが、実際の動作は、個々の基本行動の出力結果に依存する。この意味で、動作出力はボトムアップ的に決定され、さらに、この動作によって次の観測情報が得られ、次の行動調停則の行動重みが決定される。つまり、トップダウン的な行動調停則とボトムアップ的な動作生成によりロボットは構成され、さらに行動調停則と基本行動が、互いに機能を限定し合う相互依存の入れ子構造を成す。このような入れ子構造を持つ多目的行動調停に基づくロボットの学習に関して、環境条件に合わせた構造的な学習手法を提案し、動作の滑らかさや基本行動の再利用性などの議論を行う。

本論文は、次のように構成される。

第一章では、本研究を行うにあたり、従来のロボットシステムの制御機構や学習機構の背景として、行動に基づくロボティクスや進化的ロボティクス、強化学習に基づくロボットの学習について説明し、それぞれ問題点として、観測情報の時系列の軽視、学習環境への特化、要素行動の非分離などを指摘する。それらを踏まえ、多目的かつ滑らかな動作の生成と再利用できる行動獲得の実現を本研究の目的とすることを述べる。

第二章では、ロボットの適用環境の分類と、ソフトコンピューティングを中心とした学習手法について説明する。また、周辺研究についても概説し、獲得済みモジュールの再利用困難性について説明する。

第三章では、多目的行動調停に基づく移動ロボットのナビゲーション問題を例にとりあげ、静的未知環境における基本行動の獲得を通して、動作の滑らかさや、学習環境と異なる環境での行動の再利用性について議論する。具体的には、距離センサ情報を入力とし、モータの速度と操舵角を出力とするファジィコントローラにより基本行動を構築する。基

本行動には、目標追従と障害物回避、壁面併走を用い、行動調停により障害物を避けながらゴールへ向かうといった多目的な動作の実現を試みる。行動調停則にはプロダクションルールを適用し、時系列観測情報を考慮するように設定する。静的未知環境は、ロボットが動作している間、環境内が変化せず、ロボットの観測情報が局所的であることから、行動獲得に進化的計算を適用する。進化的計算は、タスクの達成度合い（例えば、移動時間、移動距離など）から一連の行動が獲得できる手法であるが、環境が複雑になれば多くの繰り返し試行が必要になる。そこで、学習初期において、各基本行動に対して理想出力を設定しデルタルールによるオンライン学習を適用する。この提案手法を計算機シミュレーションにより実験し有効性を検証する。まず、静的未知環境下で、観測情報に対して滑らかな主行動の遷移が行われる動作が獲得できることを示した。次に、複数の静的未知環境下で、反射的な行動切り替え型ロボットと比較することで、多目的行動調停を用いることで基本行動が短い世代で獲得できることを示した。さらに、学習環境とは異なる環境下において、多目的行動調停を用いることで、学習環境と同様の性能が得られることを示した。また、行動切り替え型のロボットが、学習環境に依存して個々の基本行動が特化され、学習時と異なる環境で上手く動作できないことを確認した。つまり、多目的行動調停を用いたロボットの行動獲得は、観測情報に特化するのではなく、観測情報の変化に依存してトップダウン的に与えられる役割に適応すると考えられる。そのため、多目的行動調停に基づくロボットは、学習時とは異なる静的未知環境に対しても、環境の障害物の配置や疎密さなどが大きく変わらない限り、問題なく対応できる。これらの結果から、静的未知環境に提案手法を用いることで、再利用性の高い基本行動が獲得できることが明らかになった。

第四章では、静的未知環境から、移動障害物を含む動的環境へと問題を拡張し、静的未知環境で学習したロボットの適応性に関して議論を行う。動的環境下では、ロボットが動作している間に環境が変化する。そのため、静的環境下で用いてきた行動調停則のヒューリスティクスとは異なり、移動障害物に対する行動重みの変更が遅れが生じ、移動障害物と衝突することがある。これに対し、移動障害物に対応する回避行動を基本行動として加え、新たな行動調停則を付加することで、移動障害物に対応することもできるが、どのような時に、どのように行動重みを変化させ、回避行動を用いれば良いかは、実際に移動障害物に遭遇してみなければ解らず、学習に多くの試行を費やす必要がある。そこで、新たな行動や行動調停則を導入するのではなく、静的未知環境で獲得した行動調停則を、部分的に調整することで、移動障害物に迅速に対応する手法を提案する。具体的には、移動障害物との遭遇による一連の動作を局所エピソードとして保持し、衝突時の緊急回避的な行動重みを教師値として、局所エピソードに遡って学習する手法である。計算機シミュレーションと実機移動ロボットを用いた実験により、基本行動を変えずに、行動調停則を一部更新するだけで、移動障害物を含む動的環境下でタスクを遂行できることが明らかになった。つまり、提案手法は、静的環境下で獲得した基本行動や行動調停則を、動的環境下で学習しなおすことなく再利用できる手法であるといえる。

(氏名： 能島 裕介 No.3)

第五章では、さらに人間と共生する人型のパートナーロボットの構築へ向けて、手渡し動作を例に、基本行動の獲得と多目的行動調停の適用に関して議論を行う。ロボットの行動に対する人間の評価は、事前に特定することができず、人間とロボットとの距離や、ロボットの動作パターンによって時々刻々と変化しうる。最初に、実際の評価値入力により生成される人間の評価モデルの同定と、同時に評価の高い手渡し動作を探索する対話的軌道生成手法を提案する。軌道生成には、進化的計算を適用し、手渡し動作の軌道を生成する中間姿勢の集合を個体として、人間の評価モデルを含む評価関数を用いて探索を行う。計算機シミュレーションでの軌道探索を内部評価ループで行い、一定回数の内部探索後に軌道の一つを選択し、実際にアームを動かすことで人間の評価を得る。人間の評価モデルとしてファジィ状態価値関数を適用し、人間からの評価値入力により、利益共有法により学習する。学習したファジィ状態価値関数を用いて再び内部評価ループにより評価値が高くなる軌道を探索する。これらを繰り返すことで、人間の評価モデルが同定されつつ、少ない評価回数で手渡し動作が獲得できることを示した。また、得られた軌道を用いて、ニューラルネットワークに現在の関節角度に対する次の変化量を学習させることで、手渡し動作の基本行動として獲得できることを示した。この手渡し行動と他の基本行動（腕振り行動と待機行動）の出力を多目的行動調停により融合することで、人間との距離の違いに合わせた滑らかな主行動の遷移が可能であることを示した。さらに、融合された一連の動作において、手渡し動作のための距離を学習することで、適した動作を適した距離でロボットが実行できることを示した。このように、対話により獲得した基本行動とその行動出力を簡単な調停則により融合した多目的な動作は、従来の人型ロボットにおける完全な事前作り込みを軽減することができると思われる。

第六章では、提案手法により実現できた多目的かつ滑らかな動作の生成と、獲得した行動の再利用性についてまとめる。また、静的未知環境や動的環境下での構造的な学習手法において、基本行動と行動調停則の学習をどのタイミングでどのくらい行うかという問題を今後の課題として述べる。さらに、人間との共生を考えた上で、行動を獲得し学習し続ける機能がロボットに必要となることを述べる。

氏名	能島 裕介		
論文 題目	多目的行動調停に基づく知能ロボットの行動獲得に関する研究		
審査 委員	区 分	職 名	氏 名
	主 査	教授	小島 史男
	副 査	教授	田浦 俊春
	副 査	教授	上原 邦昭
	副 査		
			印
			印

要 旨

知能ロボットがタスクを遂行するために、複数の基本行動を用いる枠組みである行動に基づくロボティクスや、タスクの達成度合いから行動を獲得する進化的ロボティクスの研究が盛んに行われているが、ほとんどの研究では動作の滑らかさや行動の再利用性について、議論されていないのが現状である。

本研究では、時系列観測情報を考慮した多目的行動調停に基づくロボットの枠組みを提案し、様々な状況に対応するための多目的かつ滑らかな動作生成及び、タスクを達成するために必要となる基本行動と行動調停則の学習方法を提案している。多目的行動調停とは、どの基本行動をどのくらい用いるかをヒューリスティックに決定し、各基本行動の行動出力を重み付け平均で融合する手法である。この行動調停則は基本行動に対してトップダウン的に役割が与えられるが、実際の動作は、個々の基本行動の推論結果に依存する。この意味で、動作出力はボトムアップ的に決定され、さらに、この動作によって次の観測情報が得られ、次の行動調停則の行動重みが決定される。つまり、トップダウン的な行動調停則とボトムアップ的な動作生成により、ロボットは構成され、さらに行動調停則と基本行動が、互いに機能を限定し合う相互依存の入れ子構造を成す。このような入れ子構造を持つ多目的行動調停に基づくロボットの学習に関して、環境条件に合わせた構造的な学習手法を提案し、動作の滑らかさや基本行動の再利用性などの議論を行っている。

本論文は、次のように構成されている。

第一章では、本研究を行うにあたり、従来のロボットシステムの制御機構や学習機構の背景として、行動に基づくロボティクスや進化的ロボティクス、強化学習に基づくロボットの学習について説明し、それぞれ問題点として、観測情報の時系列の軽視、学習環境への特化、要素行動の非分離などを指摘している。それらを踏まえて、多目的かつ滑らかな動作の生成と再利用できる行動獲得の実現を本研究の目的とすることを述べている。

第二章では、ロボットの適用環境を分類し、それぞれの環境で用いることができる情報や学習手法について説明している。また、本論文で提案している制御手法や学習手法の構成要素となるソフトコンピューティングや強化学習などを概観し、本研究で中心的な考えとなる学習機構の構造化を問題環境の特徴を考慮して行うべきであると示唆している。さらに、周辺研究における獲得済みモジュールの再利用困難性についても説明している。

第三章では、多目的行動調停に基づく移動ロボットのナビゲーション問題を例にとりあげ、静的未知環境における基本行動の獲得を通して、動作の滑らかさや学習環境とは異なる環境での基本行動の再利用性について述べている。具体的には、まず、距離センサ情報を入力とし、モータの速度と操舵角を出力とするファジィコントローラにより基本行動を構築し、各基本行動の出力を行動重みに基づいて重み付け平均し実際の動作出力を生成する。行動重みは、観測情報の変化を考慮した行動調停則により更新され、各基本行動の重要度を逐次変更する。基本行動には、目標追従と障害物回避、壁面併走を用い、障害物を避けながらゴールへ向かうといった多目的な動作の実現を試みる。静的未知環境は、ロボットが動作している間、環境内が変化せず、ロボットが観測できる情報が局所的であることから、行動獲得に進化的計算を適用している。進化的計算は、タスクの達成度合い（例えば、移動時間、移動距離など）から一連の行動が獲得できる手法であるが、環境が複雑になれば多くの繰り返し試行が必要になる。そこで、学習初期において、各基本行動に対して理想出力を設定しデルタルールによるオンライン学習を導入している。この提案手法を計算機シミュレーションにより実験し有効性を検証している。まず静的未知環境下で、観測情報に対して滑らかな主行動の遷移が行われる動作が獲得できることを示している。次に、複数の静的未知環境下で、基本行動が短い世代で獲得できることを示している。さらに、学習環境とは異なる環境下においても、学習環境と同様の性能が得られることを示している。

氏名	能島 裕介
----	-------

一方、比較として、反射的な行動切り替え型ロボットにおいても同様のシミュレーションを行い、学習環境に依存して個々の基本行動が特化され、学習時と異なる環境で上手く動作できないことを確認している。多目的行動調停を用いたロボットの行動獲得は、観測情報に特化されるのではなく、観測情報の変化に依存してトップダウン的に与えられる役割に対して適応すると考えられる。そのため、多目的行動調停に基づくロボットは、学習時とは異なる静的未知環境に対しても、環境の障害物の配置や疎密などが大きく変わらないような環境であれば、問題なく対応できることが明らかされている。つまり、行動調停則の設計の段階でのヒューリスティクスから、環境が大きく異なる限りにおいて、学習環境と同様の性能が得られ、基本行動の再利用性が高いことが示されている。

第四章では、静的未知環境から、移動障害物を含む動的環境へと問題を拡張し静的未知環境で学習したロボットの適応性に関して議論を行っている。動的環境下では、ロボットが動作している間に環境が変化する。そのため、静的環境下で用いてきた行動調停則のヒューリスティクスとは異なり、移動障害物に対する行動重みの変更に遅れが生じ衝突することがある。これに対し、移動障害物に対応する回避行動を基本行動として加え、新たな行動調停則を付加することで、移動障害物に対応することもできるが、どのような時に、どのように行動重みを変化させ、回避行動を用いれば良いかは、実際に移動障害物に遭遇してみなければならず、学習に多くの時間を費やす必要がある。そこで、新たな行動や行動調停則を導入するのではなく、静的未知環境で獲得した行動調停則を、部分的に調整することで、移動障害物に対応する手法を提案している。移動障害物との遭遇による一連の動作を局所エピソードとして保持し、衝突時の緊急回避的な行動重みを局所エピソードに遡って学習することで、移動障害物との同じような遭遇に対し、衝突を回避できることを示している。この結果から、基本行動を変えずに、トップダウン的な行動調停則を一部更新するだけで、移動障害物を含む動的環境に適応できることが明らかにされた。提案手法は、新たに基本行動や行動調停則を導入するよりも少ない学習回数で移動障害物に対応することができ、動的環境下における適応学習として有効な手法であると考えられる。

以上の静的未知環境及び動的環境の実験結果から、移動ロボットにおける多目的行動調停と環境の特徴に合わせた学習手法の有効性が示されている。

第五章では、さらに人間と共生する人型のパートナーロボットの構築へ向けて、手渡し動作を例に、基本行動の獲得と多目的行動調停の適用に関して議論を行っている。ロボットの行動に対する人間の評価基準は、事前に特定することができず、人間とロボットとの距離や、ロボットの動作パターンによって時々刻々と変化しうると考えられ、人間を含む環境は、移動障害物を含む動的環境よりも環境が持つダイナミクスが変化する問題として帰着できる。最初に、実際の評価値入力により生成される人間の評価モデルの同定と、同時に評価の高い手渡し動作を探索する対話的軌道生成手法を提案している。軌道生成には、進化的計算を適用し、手渡し動作の軌道を生成する中間姿勢の集合を個体として、人間の評価モデルを含む評価関数を用いて探索を行っている。計算機シミュレーションでの探索を内部評価とし、一定回数の探索後に軌道の一つを選択し、実際にアームを動かすことで人間の評価を得る。人間の評価モデルとしてファジィ状態価値関数を適用し、人間からの評価値入力により、利益共有法により学習する。学習したファジィ状態価値関数を用いて再びシミュレーションにより評価値が高くなる軌道を探索する。これらを繰り返すことで、人間の評価モデルが同定され、人間の評価モデルを用いない方法よりも少ない評価回数で手渡し動作が獲得できることを示している。また、得られた軌道を用いて、ニューラルネットワークに現在の関節角度に対する次の変化量を学習させることで、手渡し動作の基本行動として獲得できることを示している。この手渡し行動と他の基本行動（腕振り行動と待機行動）の出力を多目的行動調停により融合することで、人間との距離の違いに合わせた滑らかな主行動の遷移が可能であることを示している。さらに、融合された一連の動作において、手渡し動作のための距離を学習することで、適した動作を適した距離でロボットが実行できることを示している。また、対話により獲得した基本行動とその行動出力を簡単な調停則により融合した多目的な動作は、従来の人型ロボットにおける完全な事前作り込みを軽減することができる。とえられる。

第六章では、提案手法により実現できた多目的かつ滑らかな動作の生成と、獲得した行動の再利用性についてまとめている。また、静的未知環境や動的環境下での構造的な学習手法において、基本行動と行動調停則の学習をどのタイミングでどのくらい行うかという問題を今後の課題として述べている。

本論文は知能ロボットの行動獲得について、その学習手法を適用環境を考慮して研究したものであり、ロボットの制御構造や学習構造について重要な知見を得たものとして価値ある集積であると認める。よって、学位申請者の能島裕介は、博士(工学)の学位を得る資格があると認める。